DOCUMENT RESUME

ED 230 570 TM 830 251

AUTHOR Lewy, Arieh
TITLE Evaluation Standards: Comments from Israel.
PUB DATE Apr 83
NOTE 14p.; Paper presented at the Annual Meeting of the American Educational Research Association (67th, Montreal, Quebec, April 11-15, 1983).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Evaluation Methods; Evaluation Utilization; *Foreign Countries; International Studies; *Program Evaluation; *Reference Materials; *Standards
IDENTIFIERS *Israel; Joint Committee on Standards for Educ Evaluation; *Standards for Evaluation Educ Prog Proj Materials

ABSTRACT

The majority of "The Standards for Evaluations of Programs, Projects, and Materials" (SEPPM) turns out to be helpful for Israeli evaluators in checking their working plans and reviewing evaluation studies. Nevertheless, the lack of overriding rationale or of theoretical anchors for the SEPPM creates difficulties in adapting them to local situations which are not directly covered in the book. It would be useful to provide a more precise definition about what is a legitimate object of evaluation. A comprehensive statement about the foundations of the ethical standards would facilitate extrapolation to cases which are not specified in the SEPPM. Israeli evaluators are not expected to be active in promoting the implementation of their recommendation to the degree suggested by the SEPPM; the utilization of study results are mediated through an academic filter operated by the Chief Scientist in the Ministry of Education. Evaluation is usually conducted in a bureaucratic setting and therefore management procedures applicable to such a setting should be fully implemented in evaluation studies. (PN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

EVALUATION STANDARDS: COMMENTS FROM ISRAEL

Arieh Lewy
Tel-Aviv University
School of Education

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Arieh Lewy
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED230570

The Standards for Evaluations of Education Programs, Projects, and Materials consist of considerations and regulations which are supposed to guide the professional evaluator in planning and carrying out a study. The authors extracted ideas, rules, etc. from various fields of disciplined inquiry such as human relationship, group dynamics, management and administration sciences, political sciences, law, ethics, research methodology, etc., examined their relevancy to educational evaluation, and formulated standards in terms which directly touch upon matters of the evaluation profession.

The Standards may serve as a tool for evaluators to examine the quality of their work, to improve their performance, to criticize and to evaluate published studies and in case of doubt and uncertainty they may serve as guidelines for decision making. They are expected to fulfill the role of a Guide for the Perplexed, and in the field of evaluation we certainly have no lack of perplexed. Like other "Codes of Behavior" it may be useful for the evaluator to read it through, at least once. But the book is not designed for one-occasion reading, though its well-articulated structure helps the reader to remember many of the ideas contained in it. It is rather designed as a reference book to be used whenever one encounters doubts, controversies, differences of opinions, etc. On such occasions one is expected to look for relevant paragraphs, which may facilitate the solution of a particular problem.

TM 830 251

The authors view the publication of the Standards as the starting point of a long road. Additional regulations (or Standards) and amendments to the existing ones will follow as needed.

The American Bar Association published a set of standards as early as 1908 (and quite likely some agreed upon standards prevailing in the profession even before that date) and since then systematic work has been done to update the standards and inform the interested parties about new developments in the field. Lawyers and physicians obtain information about conceptions in professional ethics also through communications in mass media about relevant court decisions.

Educational evaluation has now taken its first step in establishing professional standards. One may expect that the efforts made in the USA will motivate other countries to draft evaluation standards suited to their particular context. Despite the diverse conditions surrounding evaluation activities in various countries, the American Standards seem likely to serve as a model for other countries, too. Moreover it is expected that for a long time the Joint Committee's Standards will be the only available series of standards in the field of evaluation and will serve as a guidebook for evaluators throughout the whole world. Therefore it may be useful to gather comments about the Standards from evaluators in other countries, with the aim of separating out standards which have universal validity from those which are applicable only in a particular context. Israeli evaluators have already commented on the Standards (Nevo 1982). While Nevo examined the Standards in light of the differences between the two educational systems, this paper will examine the relevance of the Standards to questions raised by evaluators in Israel.

The Rationale of the Standards

The reader who received the Standards in the form of a bound volume and has not had the chance of following the discussions which undoubtedly preceded the publication would like to obtain some orientation about the rationale, or the philosophical roots of the proposed codes. This would be useful especially for those who may be willing to examine the implications of the Standards in a milieu which is different from the one in which they were generated. Information about ideas which guided the Committee in formulating the Standards, the specification of principles which supported decision making during the process of the deliberations, and details about possibly unresolved issues could enhance the understanding of the standards and facilitate their application to changing situations. It may well be that the authors deliberately avoided taking positions on controversial issues, feeling that frequently it is easier to develop a consensus about a behavioral code than about a philosophical idea, but application and adaptation requires more than knowing the rule. Therefore some questions related to the rationale of the Standards will be presented here in the course of examining their relevance to the concerns of evaluators in Israel.

The Theoretical Basis of Evaluation Activities

Do the standards reflect a certain position about the relationship between research and evaluation? Are there methodological standards which apply to evaluation and do not apply to research? Which are these standards, if any?

The Standards present evaluation as an activity which examines the worth of action taken by others. The evaluation team is not the agent proposing changes. It is responsible for approving or improving changes suggested by others. This role division is by no means a common practice in Isael. There are cases where

the evaluators initiate the innovation, and where the innovation is the product of carefully designed development and evaluation activities in which the evaluator has the leading role, assuming responsibility for the innovation in the capacity of principal investigator.

Thus, for example, the Israeli universities initiated a tutoring project in which undergraduate students volunteered to tutor disadvantaged learners. No attempts have been made to evaluate the whole project across universities and across substrata of the target population. Neither the initiators of the project nor professional evaluators felt challenged to carry out a summative evaluation study. Instead of conducting a non-interventional summative evaluation of the whole project, a team of evaluators initiated a study designed to identify those context and treatment variables (and the interactions among them) which contributed to the success of the program, as measured by various dependent variables, such as cognitive achievements, motivation, school attitude, self-esteem, reduction of alienation, realistic goal-setting, etc. The team strove to develop an inventory of "context-activities-outcomes" configurations, which might then serve as a guide for planning tutoring activities. One should note that the study was the initiative of an evaluation team. The series of activities was carried out in line with the tradition (if there is such a thing) of evaluation practice. The direct client was not the agency which initiated the tutoring project. According to Israeli conceptions, the study was labeled an evaluation study.

Would it be viewed as such by the authors of the Standards, too?

If not, why not? How should one set limits? In the past some development studies of this type have in the USA been considered as pertaining to the domain

of evaluation. Thus, for example, studies which lead toward the concept of Mastery Learning have been considered evaluation studies. In these studies evaluators did not examine the work of programs developed by others, but themselves developed a program on basis of a series of evaluation studies. Would such an activity be considered evaluation by the authors of the Standards? Are there methodological caveats which apply to such studies and do not apply to evaluation designed to assess the merits of a program initiated by others than the evaluator himself?

Another question related to the object of evaluation, which emerged in Israel, is the specificity or uniqueness of the treatment being evaluated. The professional community in Israel would distinguish between the evaluation of treatments rooted in well-established or commonly accepted theories or in treatments which have a bearing on the verification of a given theory, on the one hand, and the evaluation of innovations or actions which are based on practical considerations, on the other. In our local context, little, if any, respect would be given to evaluation studies of the second type. In Israel the evaluator is supposed to specify the theoretical considerations guiding his work. Thus, a study which examines the impact of fund allocation on achievement scores (quoted in the Standards as an example), would hardly be considered a respectable evaluation endeavor, unless the evaluation procedures were based on some explanatory theory.

Do the Standards require activities of this type or do they postulate some theory-orientated justification of actions?

Ethical Considerations

Issues of ethics are pertinent to several aspects of evaluation studies. Problems of ethics related to evaluation and research have been dealt with extensively both in the field of education and in other branches of Behavioral and Life Sciences (Boruch and Cecil 1979; Perloff and Perloff 1980). Moreover, professional ethics have evolved into an independent field of study within the framework of Ethics and within various professions (such as law, medicine, psychiatry, advertisement, journalism, etc.). Research has strived to explore the theoretical and moral foundations of professional ethics (Goldman 1980), and court decisions which have bearings on establishing standards of professional behavior are frequently encountered.

The evaluation Standards address several issues in the domain of ethics, and lists three standards under the heading "Ethics" in the index: formal obligations, the public's right to know, and the rights of human subjects. In fact other standards can be identified as reflecting ethical principles. The standards related to conflict of interest and decisions about accepting a contract also belong in this category. Researchers in this country encountered problems with an ethical basis in addition to those treated in the Standards. Thus for example the codes related to competition for contracts and advertisement, which receive extensive treatment in the ethical Standards of the legal and medical professions, are pertinent to educational evaluation as well.

The authors list ethical standards in operational terms. Such a presentation facilitates their utilization in concrete situations, but fails to guide a person in situations which are not specified in detail by the standards. Goldman (1980) raises the question of whether the ethical codes for various professions can be

derived from general principles of ethics. Clearly enough the dishonest use of professional status for private gain, or the violation of contractual agreements need not be regulated by strongly differentiated professional codes; the ordinary moral categories are sufficient to assess such conduct. But some professions may have codes of behavior which cannot be derived from general moral categories. The judge's sentence should be based on the application of the law. He is not allowed to make decisions based on his moral conviction. Lawyers are expected not to publicize circumstances which may aggravate the situation of their defendent, provided their failure to do so does not cause harm to others. What is the status of the evaluators? Are there moral obligations for an evaluator which cannot be derived from general ethical codes? If so, what are they?

Without relying on some comprehensive general principles one can hardly further develop ethical standards let alone adapt them to changing circumstances.

Utility Perceptions

The Standards emphasize three categories of behavior which determine the social utility of an evaluation study. First one should produce results which have bearings on decisions to be taken. Secondly, results should be accessable to the decision maker and thirdly the evaluator should be active in promoting the implementation of his recommendations. All three of these categories are important also in Israel, but probably the third one has a more limited role here than it has in the USA.

One should note that the complex relationship between making a professionally sound recommendation and motivating people to act according to these recommendations is not unique to the field of educational evaluation (Lewy 1980).

In Israel the channel of communication between the educational authorities and the academic community is obviously more direct and more efficient than it appears to be in the USA. Since the Israeli educational system is a highly centralized one, most pervasive decisions touching upon educational practice are made at the national level by the Ministry of Education and Culture. One of the senior positions in the Ministry of Education, as well as in other Israeli ministries, is that of Chief Scientist. The Chief Scientist is a member of the academic community and holds a part-time position in the Ministry. His major role is to serve as a link between the educational administration and the academic community. He is in charge of commissioning research studies, examining their quality and assisting senior administrators to implement research recommendations. The onus of implementing recommendations is transferred from the researcher or the evaluator to theChief Scientist of the Ministry (Kugelmass 1981). Since the implementation is mediated through the Chief Scientist and his staff, it turns out that the scientific qualities of the study have a greater weight than its communicative qualities in determining the chances of utilization. Rather than a disregard of research findings by administrators, in the past we encountered cases where administrators were eager to act on basis of research finding which were not yet fully validated.

The Standards encourage the communication of the interim results of studies to the client and perceive such communication as a tool for promoting the utilization of the results. The Standards do caution against the premature disclosure of findings which are not fully validated, but in general seem to favor continuous communication with the client about interim results.

A recent case of a semi-official release of the interim results of a study raised doubts about the desirability of such action, and created an awareness of the need for establishing more stringent standards as to what, when, and under what circumstances information should be, and may be communicated to the client. The content of the above mentioned semi-official release of information generated a vehement public reaction in the newspapers. While politicians and the mass media demanded action on basis of the interim results, the academic community was embarrassed because they could not comment on the validity of the results.

The Standards devote a great deal of attention to cases like the one noted. A caveat states "Do not confuse full and frank disclosure with premature disclosure of information (p. 75)." Another standard suggests that data should be made available for checking by responsible reviewers (p. 108). Nevertheless, the controversy which emerged around the Israeli case described above called attention to the need of providing more specific guidelines about the following questions:

1. When is it desirable to communicate the interim results of a study, and how?
2. Are there cases when an evaluation team is supposed to act like a judiciary investigation committee or a jury which is barred from disclosing details about the process of deliberation before a final decision has been reached?
3. What are the obligations of the evaluator if news media present an unbalanced selection of the totality of findings included in a report or released to the public?

Management Procedures

The standards provide a set of suggestions which have their roots in public administration and business management practices. The fact that such standards should be specified suggest that academic people trained in educational evaluation are inclined to disregard such mundane obligations as "maintaining accurate records of funding and expenditures." Here again are standards which are not unique to the field of evaluation. They reflect general standards applying to situations of various types involving money which is not private property.

Some management standards are more characteristic for evaluation activities, thus for example, the Standards suggest maintaining a log of unusual events (p. 105), and preserving and making raw data available for responsibly planned reviews (p. 108).

Evaluation is usually conducted in a bureaucratic setting and therefore regulations applicable to such a setting should be fully implemented in evaluation studies too. Israeli evaluators (and quite likely evaluators in other countries) need to be reminded of the existence of such standards and of their pertinence to evaluation practice. In this context it should be noted that the word "standard" has several meanings. The Oxford Dictionary defines "standard" as a degree of quality viewed as a measure of what is adequate for some purpose "or as, a definite level of excellence." Ethical and methodological standards reflect these meaning of the word standard. But the word also has an entirely different meaning. It may signify an arbitrarily agreed form or shape of a certain phenomenon or object. Thus for example, electric plugs have a standard size. In the absence of agreement on a standard size for electric plugs, it would be very inconvenient to use electric appliances.

Bureaucratic management systems, too, develop standard operating procedures, arbitrarily setting certain acceptable operational rules. Due to the existence of such standards, a well trained accountant can come to any business firm and audit the books. Thus outside control can be exercised on the basis of inspecting the available written documents, without any need to ask for explanations of secret signs or cryptic records.

In educational evaluation we are far from having standards of this type. The Joint Committee emphasized the desirability of maintaining order in the records of a study, but for the benefit of the evaluators, it would be worthwhile to appoint a sub-committee to propose more mundane regulations than those contained in the evaluation Standards. An attempt to specify types of records which should be compiled and preserved within the framework of an evaluation study is described elsewhere (Lewy 1982).

Concluding Remark

The Standards provide useful guidelines for evaluators in Israel as well as the USA. Nevertheless, the evaluator is likely to encounter many problems for which no solutions are provided in thepublished set of standards. In such cases the ancient parable about the king and his coachman is a useful guide.

The parable tells that a kingdom contained two districts separated from each other by a mountain, and accessible to each other through a single road. Sixteen feet wide, the road was flanked on one side by steep rocks and on the other by a deep precipice. The king had to ride through this road each day, and when seeking a new coachman he asked each applicant how close to the dangerous precipice he could drive the royal coach without tumbling over. The first applicant thought he could drive within two feet of the edge, the next thought that he could drive within four feet, the third said within six feet, and so on. Finally one applicant said: "If I had the honor of driving

your majesty's coach, I would keep as far away from the edge as I possibly could". And so should every evaluator. *

* (Adapted from R.L. Wise: Legal Ethics (2nd edition) Matthew Bender, New York, 1970).

References

Boruch, R. and Cecil, J. Assuring Confidentiality in Social Research Data. University of Pennsylvania Press , Philadelphia, 1979.

Goldman, Alan H. The Moral Foundations of Professional Ethics Rowman and Littlefield, Totowa NY. 1980

Kugelman, S. Considerations toward a Policy of Evaluation Research: The Case of the Chief Scientists at the Israeli Ministry of Education. Studies in Educational Evaluation Vol. 7, 1981, 161-171.

Lewy, A. Professionals and their Professions PhiDelta Kappa CEDR Quarterly Vol. 13(1), 1980, p. 1, 6.

Lewy, A. Standards for Recording and Maintaining Research Data. Scottish Educational Review V. 14, 1982, 31-37.

Nevo, D.: Applying the Evaluation Standards in a Different Social Setting. Paper presented at the 20th Congress of the International Association of Applied Psychology, Edinburgh, Scotland July 25, 1982.

Perloff R. and Perloff E.: Values, Ethics and Standards in Program Evaluation. Jossey Bass, San Francisco 1980.